Viviane Reding, three-time European Commissioner, is a force of nature. Evidence? Well, it wouldn’t be quite true to say that she single-handedly brought the right to be forgotten to the forefront of debate about privacy and data protection, but it’s not far off. Before Reding took up her post as Justice and Rights Commissioner (after a stint as Information Society Commissioner), the right to be forgotten was generally perceived as a minor sport for Luddites, privacy nuts, and those who didn’t understand what the Internet involved. Google Trends shows hardly anyone searching for the term prior to 2012, while about half the papers on Google Scholar on the topic have been published since 2013.

Reding’s muscular speeches advocating a right to be forgotten for Europeans kick-started a ruckus that has pitched the European Union (EU) against the US and privacy activists against Big Data advocates. This issue gained momentum in May 2014, when an appeal by Google Spain against a decision of the Spanish data protection authority (DPA), la Agencia Española de Protección de Datos (AEPD), was rejected by the Court of Justice of the European Union (CJEU), thereby enshrining the right to be forgotten in law. Academics, lawyers, politicians, and businesspeople have thundered about “the biggest threat to free speech on the Internet in the coming decade,” while Google’s advisory council said on the other hand that the right doesn’t exist, and Swedish Pirate Party founder Rick Falkvende went so far as to argue that it doesn’t even protect privacy. Focusing on pragmatics rather than morals, Jimmy Wales (Wikipedia founder and another member of the Council) thinks it’s “well-meaning but incoherent,” while the Economist newspaper worried that “not just giants like Google and Facebook but also innovative startups will be weighed down,” and that it “will push the Internet further towards fragmentation.”

Yet even the mighty Google is playing nicely: after a fearsome initial assault on what Google’s Global Privacy Counsel Peter Fleischer called “foggy thinking,” it is meeting information commissioners regularly and removing hundreds of thousands of URLs from its search results.

The trouble with human rights is that they affect people in their day-to-day, social, and economic existence, both in terms of the rights we can claim, and the duties they subtend. In the digital world, they affect those technologists who are architects of that world, and all its denizens, and yet discussion about them is monopolized by lawyers. This is unsurprising, as ultimately claims to rights will be adjudicated in courts. Yet rights have meanings and significance that go beyond courtroom exegesis, and we digital citizens deserve our say, too.

I don’t pretend to speak for every digital citizen, but the repercussions of this remarkably deep ruling deserve a closer look. There are many aspects of the ruling I would strongly defend. Many aspects of the way the ruling was reached are less defensible. Not being a legal scholar, I have no idea how problematic this example of the similarity between laws and sausages is (you should not watch either being made). It doesn’t leave me with a comfortable feeling, and a better means of protecting privacy rights across jurisdictions is sorely needed.

Just the Facts, Ma’am

The facts of the case are remarkably simple, yet the judgment was still unexpected. In 1998, a gentleman’s house in Spain was auctioned to cover some social security debts; this was reported in...
The Right to Be Forgotten: The Good, the Bad, and the Ugly

Key Concepts of Data Protection

The following are the basic definitions underscoring the European Union's Data Protection Directive of 1995.

- Data quality. Any processing of personal data must be lawful and fair to the individuals concerned; ... the data must be adequate, relevant, and not excessive in relation to the purposes for which they're processed ... such purposes must be explicit and legitimate and must be determined at the time of collection of the data ... the purposes of processing further to collection shall not be incompatible with the purposes as they were originally specified.

- Data processing. Processing of personal data (“processing”) shall mean any operation or set of operations which is performed upon personal data, whether or not by automatic means — such as collection, recording, organization, storage, adaptation — or alteration, retrieval, consultation, use, disclosure by transmission, dissemination, or otherwise making available, alignment or combination, blocking, erasure, or destruction.

- Data controller. Controller shall mean the natural or legal person, public authority, agency or any other body which alone or jointly with others determines the purposes and means of the processing of personal data.

- Establishment. The processing [falls under the Directive if it] is carried out in the context of the activities of an establishment of the controller on the territory of the Member State; when the same controller is established on the territory of several Member States, he must take the necessary measures to ensure that each of these establishments complies with the obligations laid down by the national law applicable.

For more information on these definitions, see http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:31995L0046.

a newspaper, partly to publicize the auction and maximize the receipts. Later, a digital archive version of the paper appeared online, and in Spain googling the gentleman’s name would reliably cause this affair to resurface.

Our hero asked the newspaper to take down the archived piece, and Google Spain to de-list it, on the grounds that the information was irrelevant and excessive (these are key concepts from the EU’s Data Protection Directive of 1995; see http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:31995L0046 and the related sidebar); he had moved on since the incident, and it was prejudicial to him and his business to have the information constantly out there.

Being nice got him nowhere, so he took his case to the AEPD, which — in contrast to other national DPAs — had been pursuing the rights of Spanish citizens to be forgotten under the Directive. The AEPD had already taken up several dozen similar cases since 2007. It backed him (against Google, though not against the newspaper), and Google Spain took up its right to challenge the judgment in the CJEU.

Google Spain’s defense relied on four separate contentions. First, search isn’t data processing; it involves locating, indexing, and even temporarily storing data, but not processing it. Second, the EU Directive didn’t apply. The search engine was run from the US by Google Inc. Third, neither Google Spain nor Google Inc. are data controllers, which the Directive regulates. Fourth, the information was lawfully published and so there could be no right to erase it.

Europe’s other DPAs seemed uninterested in the AEPD’s cruzada, while Google Spain was backed to an extent by the CJEU’s own special advisor, the Advocate General,9 leading one commentator to wonder whether it was time to forget the right to be forgotten.10

The court scotched such talk — for the first time, it established a right in this space by upholding (and arguably extending) the AEPD’s original judgment. It decided that search was data processing under the (rather wide) definition in the Directive — the data are collected, stored, retrieved, disclosed, and so on. Even so, the processing happens in the US by Google Inc. — what has that got to do with Google Spain? Plenty, said the court. First, the court argued that Google Spain was an establishment in the EU (nobody disputed this), and so Spanish law applied to it. It then went on to argue that the processing (in the US) was carried out in the context of the activities of Google Spain on the territory of the member-state Spain that were “intended to promote and sell ... advertising space offered by the search engine, which serves to make the service offered by that engine profitable.” Those advertising activities created a link between Google Spain and the search engine’s data processing; the court also argued that the Directive is meant to cover the data protection rights of EU citizens within the EU, and so it was bound to interpret the various concepts widely.

Most controversially, the court rejected Google’s third claim that it isn’t a controller. This is a responsible position, carrying with it stringent data protection responsibilities; it follows that this is a key part of the judgment. Google argued that, even if it processes personal data, it makes no distinction between personal and non-personal data, which arrive at its door in a haphazard and random
The Digital Citizen

way. It’s a passive intermediary, has no relationship with the data or the webmasters publishing it, and has no significant control over the content. The Advocate General agreed — to be a controller, “the data processing must appear to him as processing of personal data, that is ‘information relating to an identified or identifiable natural person’ in some semantically relevant way and not a mere computer code.”

But the court demurred. The search engine “determined the purposes and means of processing” within the context of the activities of Google Spain. This processing is separate from that performed by the third-party webmasters, and consists in creating “a structured overview of the information” relating to the individual searched for, which couldn’t be created in the absence of the search engine. Once more, it felt that full data protection for EU subjects could only be provided if the definition was interpreted with a wide scope.

Finally, the court decided that Google had no responsibility to contact third-party webmasters to tell them something had been de-listed, and that if the information was lawfully published (and therefore true), the information shouldn’t be removed from the Internet. The key privacy invasion is the possibility of creating an overview about an individual; the information objected to should only be de-listed therefore if the search’s keywords are the individual’s name. My privacy is contravened far more when someone searching for “Kieron O’Hara” finds that I committed some minor but embarrassing misdemeanor, than when she searches for the misdemeanor and finds my name among the perpetrators, because in the former case she’s clearly interested in me personally, whereas in the latter she’s not. There are also defenses when the individual involved is a public figure, whose private life may be of legitimate public interest.

The Good

As noted, the judgment received much flak from corporate types, free speech fundamentalists, Internet anarchistas, and privacy advocates. The propositions that these diverse people agree on are that the Directive is outdated, and its application to the networked world is dodgy; given this imperfect position, an imperfect judgment would necessarily follow.

Well, maybe. To be perfectly honest, I think there’s much to admire in the judgment, even if it has problems. Indeed, I think it’s a remarkably subtle piece of reasoning. And the problems are caused less by the actual judgment than its context. So, what’s so great about it?

First, its effects on free speech are proportionate. The information remains online, and can be googled if someone is looking for it specifically (as long as they aren’t just fishing for general information about the individual). Contrary to much hyperbole, history isn’t being changed, and nobody controls the past (actually, Google now has somewhat less control). There’s a public interest defense. The only free speech curtailed is that of a search engine to say that a certain webpage is the nth most relevant page about a particular named individual. That won’t satisfy free-speech fundamentalists, but it satisfies me.

Second, whereas the right to be forgotten could and does mean a number of things, and could — if pushed — be a protection for the rich and powerful to erase truths they don’t like, the Google Spain decision (though it has unpredictable ramifications) is a concrete expression of what can be done to make it harder (but not impossible) to access information. Privacy isn’t all or nothing.

There’s a big difference between information that’s public, and a dossier of the same information gathered together in one place. So the right to be forgotten is “only” a right to be de-listed? In the absence of arguments to say that it should be much more, good!

Third, it recognizes what we often forget, that Google isn’t the Web; rather, like Wikipedia, it’s a starting point. Someone with serious journalistic purposes will not, and should not, be deterred from investigation, although it may require greater resources than simple search. But then someone with serious purposes will know what he’s looking for, and shouldn’t be just fishing. The truth is not the set of webpages connected with a person by an algorithm, but a complex construct requiring intellectual investment to recover. For instance, many cases championed by the AEPD were sparked by newspaper archives reporting someone being charged with or convicted of an offence, without subsequently reporting acquittals or successful appeals. So the reports are, in one sense, the truth — but not the whole truth. Google does not — cannot — provide that.

Fourth, it rejects the idea that an algorithm, created by a corporation, is a neutral reflection of the state of the Internet. As I argued in my last column, algorithms have power and change our lives, and it shouldn’t be acceptable simply to present us with a black box and assure us that it’s OK.

Fifth, the judgment has found a reasonable balance between the interests of Europeans to be protected by European law, and the interests of non-Europeans. The judgment is being applied to national domains within the EU, .es, .uk, .fr, .de, and so on, and not to other domains like .com. It’s relatively easy for someone in the EU to use Google.com, and so critics have complained about the lack of protection. However, Google has an enormous slice of search traffic in Europe, and most people are initially directed to the national domain. This default setting means that getting to Google.com is an obstacle; it’s not much of an obstacle,
but it weeds out a large number of speculative searches, without hindering a serious, interested inquiry. And that’s okay — European law protects people in Europe, while other jurisdictions are unaffected. There are issues about protecting rights – we’ll come to them later – and the jurisdictional issues online are complex as we know, but this is surely reasonable.

This leads us to the sixth point, which is that the judgment restores some of the practical obscurity that protected our privacy in the past. In the days of paper and filing cabinets, information was often available, even public, but hard to get. This protected our privacy quite well – read Dickens’ *Bleak House* for a story about how hard it could be to find embarrassing or damaging information, even in the public domain. Such protection isn’t censorship, and isn’t targeted at particular types of information or data subjects. It’s accidental, random, complex, and unpredictable – which is what makes it effective.

Seventh, it allows the poor data subject some small measure of control of the way he or she appears to the outside world — what the Germans call *informational self-determination* and what the rest of us call dignity. This can really matter — for instance, rehabilitation of offenders is often supported by a legal right to suppress, in certain situations, details of spent convictions, which is severely undermined by unrestrained search.

Eighth, it requires that those who wish to be de-listed provide a good reason, showing that the information is excessive, outdated, or otherwise misleading, and allows for their wish to be countermanded by better reasons (such as the individual’s public profile). That seems absolutely correct. In particular, subjects’ rights outweigh the economic interests of the search company, which creates tension with the US focus on ensuring that privacy-respecting restrictions on information flow have the least possible effects on economic activity.14

Ninth, if life is indeed complicated for corporations, how terrible is that? In the Google Spain case, the Advocate General (a senior advisor to the court) opined that if de-listing must be decided on a case-by-case basis, the subsequent overhead on a search engine could overwhelm its ability to consider the merits of cases properly. Is that really true? The numbers of de-listing requests are big, but they’re probably outnumbered by copyright requests, and certainly by requests to delete link farmers. Like the picture of Dorian Gray, the elegant PageRank algorithm has long since been deformed by the need to weed out illegal content and spam, and this will just be another wrinkle added to its countenance, while the face of Google itself remains unchanged, returning results speedily and effectively. To foster trust it implies the completeness of its indexing, yet of course that desideratum is constantly overridden by commercial, legal, and other needs. Google can cope: it wants to curate all human knowledge and has photographed the entire world, after all. The CJEU seems to understand that better than its critics.

Tenth, information quality is an acknowledged concern in data protection rights, and therefore is a factor in deciding whether access should be restricted. The court has further recognized that information has a life cycle and that the passage of time affects its quality.14

Eleventh, the strong line taken by the court has of necessity stiffened the spine of the EU generally in its dealings with Big Technology. For instance, the European Commission is finally facing the fact that its “safe harbor” arrangement with the US15 doesn’t always ensure that the terms of the Directive are respected across the pond.15

Last but not least, this judgment is a victory for the rule of law over technological determinism and corporate power. It makes life complicated for the big corps – but that’s just the way it must be sometimes. That’s not to say that the Directive, particularly in our networked age, is good law, only that law is better than surrendering to the demands of innovation.

This is a formidable list, which shouldn’t be disregarded in the chorus of boos. What’s not to like?

**The Bad**

Well, there’s no doubt that the judgment creates problems and sets some unresolved conundrums. Presently, there’s no requirement to inform third-party publishers that their content has been de-listed. Google provides a form for disgruntled individuals to make their objections, but there’s no institutional mechanism for the publisher to provide counter-evidence (which indeed might cast the situation in a new light). On the other hand, we would not want to inform ill-intentioned publishers (revenge porn sites, for instance).

Google’s expertise, as mentioned, is unrivaled, and despite protestations, it will be able to cope. However, the judgment raises the barriers to entry in the search market, reducing competition. And smaller outfits with interests in search — archives, for instance — will have reservations about where the judgment leaves them. Only an organization with hundreds of developers, thousands of lawyers, and millions of dollars will be able to take on this kind of responsibility.

For the individual, the current mechanism depends on him or her identifying particular URLs containing excessive content. However, this doesn’t deal straightforwardly with the problem of the same content...
being copied and distributed from several pages. The wronged individual has to find all the URLs, with no proactivity assumed by the search engine. It’s still a lot of work for the underdog.

These three problems highlight a fourth – that the search engine is being put in a position where it’s becoming the judge of whether information is worthy of suppression from its search outcomes (a role it certainly doesn’t want to occupy). Google currently de-lists about 40 percent of requested URLs. There’s a right of appeal to the national DPA for individuals in the event of it ignoring their request (and if Google really wanted to stop playing ball, it could refer all such cases to DPAs, which would soon be overwhelmed), but these decisions about freedom of speech and protection of individual privacy are in almost all cases being made by a private company, which is hardly a tempting long-term scenario, especially when they are not very transparent either.

I don’t have a simple solution to these issues – they’re very troublesome. Nevertheless, I am more concerned with the unfortunate process of making the judgement, rather than its immediate ramifications.

**The Ugly**

What really stands out in this imbroglio is the role of the European Court. Its stance has followed a consistent trend of interpreting data protection rules widely and asserting its authority against tech giants. Europe may be in economic decline, says the court, but it’s still a substantial market; it may only be 7 percent of the world, but contains 19 percent of Internet users. It may be lagging in creativity and entrepreneurship, but it’s a thought leader in data protection. The court is essentially saying we’ll stand our ground, and we can’t be ignored. The court ignored the advice of its own Advocate General, and clearly endorsed what was seen as a maverick Spanish stance.

Over the last few years, it has consistently interpreted the terms of the Directive as broadly as it can, in line with the Directive, enshrined in the national law of each EU nation, which insists that “data-processing systems are designed to serve man” (in the sense of humankind) and “respect their fundamental rights and freedoms, notably the right to privacy.” For its understanding of human rights, it has taken to referring not to the European Convention for Human Rights (ECHR), whose language is relatively woolly and which only protects privacy, but to the more recent Charter of Fundamental Rights of the European Union, which took legal effect only in 2009, and which includes data protection rights alongside those of privacy. At the same time, the court has also interpreted some of the limitations and caveats in the Directive (for example, exemptions for journalistic purposes) in a remarkably narrow way.

The court focuses on the Directive’s aims for a high level of protection of fundamental rights. Yet the Directive itself, which predates the Charter, only refers to the ECHR (for instance, in recital 10 of the Directive’s preamble, explicitly referred to in the Google Spain judgment), and so the court’s reliance on the Charter to achieve the Directive’s aims is somewhat paradoxical to the lay observer. The Directive was written in a world where data protection was a new(ish) concept, and so may have performed has aggressively been drafted – but in combination with the Charter, which itself aggressively insists on the data protection rights of EU citizens, seems to have become a powerful tool. It is not absurd to imagine data protection, rather than the more traditional libel suit, becoming the technique à la mode for the powerful to resist pesky journalists, campaigning groups, and concerned citizens.

Privileging privacy and bypassing the ECHR is in line with other recent judgments, for example the extraordinary outcome of the so-called Bavarian Lager case (http://ec.europa.eu/dgs/legal_service/arrets/08c028_en.pdf), which prevented the release of names of people attending an official meeting with the European Commission, on the grounds that to do so would negatively impact their privacy. A lower court had argued that, as this was a business meeting with governmental representatives, there were no negative impacts on the right of these people to a private life. The European Court annulled that argument – taking the key definition of privacy from a specific regulation implementing the data protection regime, not the Convention of Human Rights.

Arcane? Sure. The upshot is that privacy trumps transparency, even when we’re clearly outside the individual’s private life.

The result is a position where the court is able to place strong constraints on those handling information. The position of data controller is a responsible one – it assumes a direct connection with the data, and decision-making power over it – but its demarcation is becoming less certain. Google’s connection with the personal data it deals with is tenuous – the data that it handles arrives randomly and it makes no attempt to distinguish personal data from the non-personal. On the other hand, if Google weren’t counted as a data controller, then the court would have no power to make Google Spain do anything, or to protect EU citizens from this particular harm. Has the court staked out an activist position for itself by means of its interpretations of the various terms and definitions of data protection?

If so, how will this impact other data handling organizations?
presumably affected by this ruling? The judgment’s superficially narrow scope will ramify across all sorts of tech companies and information crunchers, and the strong requirements on data controllers may have serious effects on the way the Internet works, because it doesn’t seem that all such companies could possibly be compliant. At the moment, we just don’t know how that will play out, partly because the DPAs are extremely bad at enforcing data protection law, as evinced by the perception of data protection in Europe as a “box-ticking” exercise, and widespread non-compliance even within the EU itself. This may be the kind of law that’s better not enforced.

The court seems to determine its own powers as it goes on, choosing narrow or wide interpretations of terms as suits its activist mission. Its independence is admirable; its autonomy perhaps less so. A European legal conversation is happening, and there’s a transatlantic debate that may well become more vocal when the new Data Protection Regulation starts to emerge from its sheltered chrysalis. Further questions also nag.

- What about individuals and companies in other countries? This is often painted as an EU versus US contest, and the rest of the world must feel a bit like shuttlecocks in a game of badminton. Will they feel forced to follow Europe’s “moral leadership,” even if they don’t agree? Could the EU’s approach work if it was internationally isolated?
- If more people and companies become classed as data controllers, will they start to demand rights to go with the onerous responsibilities — for example, free speech rights?
- Can underfunded European DPAs keep up with the workload created by an activist court?

On the other hand, the court has served notice of its intentions; maybe this is an opportunity for dialogue to begin across the different jurisdictions. It’s quite clear that there are other fights that the CJEU could pick if it wished. For instance, the regime of notice-and-consent is virtually broken, as beleaguered individuals can only cope with companies’ unreadable privacy policies and the complexity of the flow of personal data by ignoring them. It will take much work to improve consent (a job worth doing), but the CJEU has maneuvered itself into a position where it could make serious inroads into the business models of data gatherers who rely on data subjects’ uncritical consent. Similarly, Big Data drives a coach and horses through the data protection principle of use limitation, and if the right case came before it, the CJEU could deal a big blow to the new paradigm.

It has long been blithely assumed that the societal benefits of data processing will be preserved by revising data protection principles, and suggestions have been made on how to do that. However, so stately is EU lawmaking that the new Data Protection Regulation isn’t going to appear soon, and when it does it’s likely to follow the direction of the CJEU’s jurisprudence. It’s plausible that the business models of Big Data, profiling, advertising-and-surveillance, and social networking companies will be damaged in Europe. It’s time for some serious international discussion at a high level about what business practices make sense for companies and privacy-aware individuals alike. I suspect that user-centric personal data management tools will be an important part of the equation, but that’s a topic for another day. References


Acknowledgments

This work is supported under SOCIAM: The Theory and Practice of Social Machines, funded by the UK Engineering and Physical Sciences Research Council (EPSRC) under grant EP/J017728/1.

Kieron O’Hara is a senior research fellow in the Web and Internet Science Group in the Electronics and Computer Science Department at the University of Southampton. His research interests include trust, privacy, open data, and Web science. O’Hara has a DPhil in philosophy from the University of Oxford. Contact him at kmo@ecs.soton.ac.uk.

Selected CS articles and columns are also available for free at http://ComputingNow.computer.org.